NASA Technical Memorandum 88254

Crew Communications as a Factor in Aviation Accidents

Joseph Goguen, Structural Semantics, Palo Alto, California and SRI International, Menlo Park, California
Charlotte Linde, Structural Semantics, Palo Alto, California
Miles Murphy, Ames Research Center, Moffett Field, California

August 1986

NASA
National Aeronautics and Space Administration

Ames Research Center
Moffett Field, California 94035

CREW COMMUNICATION AS A FACTOR IN AVIATION ACCIDENTS

Joseph Goguen,* Charlotte Linde,† and Miles Murphy

Ames Research Center

SUMMARY

A method for the detailed analysis of within-crew communications is developed and applied in formulating and testing several hypotheses about the basic structure of the aircrew communication process. Planning and explanation are shown to be well-structured discourse types, described by formal rules. These formal rules are integrated with those describing the other most important discourse type within the cockpit: the command-and-control speech act chain. Command-and-control discourse is described as a sequence of speech acts for making requests (including orders and suggestions), for making reports, for supporting or challenging statements, and for acknowledging previous speech acts. Mitigation level, a linguistic indication of indirectness and tentativeness in speech, was an important variable in several hypotheses. Testing these hypotheses showed that the speech of subordinates is more mitigated than the speech of superiors, that the speech of all crewmembers is less mitigated when they know that they are in either a problem or emergency situation, and that mitigation is a factor in failures of crewmembers to initiate discussion of new topics or have suggestions ratified by the captain. The test results also indicated that planning and explanation are more frequently performed by captains than by other crewmembers, are done more during crew-recognized problems, and are done less during crew-recognized emergencies.

INTRODUCTION

The motivation for the study reported here is to reduce the incidence of those air transport accidents caused wholly or in part by problems in crew communication and coordination. A major objective is to determine those communication patterns which actually are most effective in specific situations; this requires development of methods for assessing the effectiveness of crew communication patterns. It is hoped that these results will lead to the development of new methods for training crews to communicate more effectively, and will provide guidelines for the design of aviation procedures and equipment.

*Structural Semantics, P.O. Box 707, Palo Alto, CA 94302, and SRI International, Menlo Park, CA 94025.

†Structural Semantics, P.O. Box 707, Palo Alto, CA 94302.

The two main contributions of this study are a set of hypotheses about aircrew communication patterns and the development of a novel methodology for formulating and testing such hypotheses. Transcripts from eight commercial aviation accidents were used as data. The second section presents the procedures used, including definitions for the variables occurring in the hypotheses, and the third section gives a precise treatment of the results. Examples from the United/Portland/78 transcript are used as a source for our examples because of their relevance to our research topic and their familiarity to the aviation community (NTSB, 1979).

This methodology is novel in its use of linguistic investigation of actual crew discussions and its use of aviation accident transcripts as data, which yielded an empirically grounded formal description of communication patterns in the cockpit. The following subsection discusses the linguistic background of the study.

Linguistic Theories

To facilitate our description of cockpit communication, we have adapted or created a number of linguistic theories. These include speech act theory and formal theories for certain discourse types: planning, explanation, and command-and-control. These theories support the linguistic variables used in our hypotheses. The variables include mitigation/aggravation level, crew-recognized emergency (CRE), crew-recognized problem (CRP), operational relevance, topic success or failure, and draft order and ratification. We now turn to a brief discussion of these underlying linguistic theories.

<u>Speech Act Theory</u>- Speech act theory is a view of language which focusses on the operational aspect of sentences. A speech act is a sentence which achieves some effect in the world. We call this the "social force" of the speech act. The fundamental insight of speech act theory is that some sentences, such as (1) <u>describe</u> or <u>report</u> a state of the world, whereas other sentences, such as (2) <u>create</u> a state of the world.

(1) There's a thunderstorm ahead.
(2) I declare this bridge open.

These examples express their social force directly. However, there are also speech acts which express their most probable social force indirectly, by using a linguistic form which is not to be interpreted literally. Examples from the United/Portland/78 transcript are indicated by giving the time of the utterance in parentheses. This transcript is used in our examples because of its relevance to our research topic and its familiarity to the aviation community (they are identified by the cockpit-area microphone noise of sound source (CAM)). For example

(3) CAM-1 What I need is the wind, really (1755:13)

is literally an <u>expressive</u>, in which the captain expresses his psychological state of "needing" information about the wind. However, given the context in which it was spoken, its social force might be given as the directive

(4) Give me the wind.

The basic question about indirect speech acts is how it is possible that one speech act is interpreted as another. To answer this question, speech act theory (Searle, 1979) uses "felicity conditions," which are conditions that must be satisfied for a speech act of a given kind to be uttered "felicitously" (also termed "nondefectively"). These conditions include preparatory conditions, propositional content conditions, sincerity conditions, an essential condition, and possibly others. "Preparatory conditions" include the conditions that must be satisfied before the utterance is made; for example, for an order that the speaker must have appropriate authority over the addressee and that the addressee is able to perform the act; or for a promise, that it is not obvious that what is promised would otherwise occur. "Propositional content conditions" express constraints on the propositional content; for example, for a promise that it express a future act by the speaker. "Sincerity conditions" concern the speaker's internal states, including his or her intentions. For example, in a request that the addressee perform an act A, the speaker should really want the addressee to do A. The "essential condition" defines the desired effect of the speech act upon the addressee.

The most obvious way to accomplish a speech act indirectly is to make reference to one of its felicity conditions. For example, a sincerity condition for a request that the addressee make a report is that the speaker should really want to know the contents of this report. Considerable literature on indirect speech acts is available in the fields of linguistics, philosophy of language, artificial intelligence, and psychology (e.g., Searle, 1979; Gordon and Lakoff, 1971; Gazdar, 1979; Labov and Fanshel, 1977). Our discussion summarizes the approach of Searle (1979) which underlies most other approaches. Speech act theory also provides a system of classification of possible types of speech acts. We have modified this system to provide an inclusive listing of the speech acts found in cockpit communication. These are "requests," including orders, requests, suggestions, and questions; "reports," "declarations," and "acknowledgments." The set of all speech acts in the eight transcripts described in the section "Sampling Procedures" constitute the basic sample on which our hypotheses were tested.

<u>The Discourse Unit</u>- Although speech act theory is of great value to the study of crew communication, our account would be quite incomplete if it remained at the level of the simple sentence. It is necessary to study larger units as well. The larger unit of language that we have found appropriate for this study is called the discourse unit. A "discourse unit" is a segment of spoken language, longer than a single sentence, having initial and final boundaries that are socially recognizable, and having a formally definable internal structure. (This definition generalizes the criteria (given by Labov, 1972) for the narrative of personal experience.) Discourse types that have been studied include the narrative, the spatial description (Linde, 1974; Linde and Labov, 1975), the joke (Sachs, 1974), small group

planning (Linde and Goguen, 1978), and explanation (Goguen, Weiner, and Linde, 1983). There are a number of points to be made about the definition of discourse unit:

1. Level of unit. In the linguistic hierarchy, the discourse unit is immediately above the sentence, and hence is composed of sentences.

2. Socially recognized boundaries. The discourse unit has boundaries which are recognized as such by the participants in the conversation. These boundaries are often recognized through their stereotyped form; for example, "They lived happily ever after" as the end of a fairy tale; "It seems there was a . . ." as the beginning of a joke; or "And that was it" as the end of a narrative. In other cases, they may be recognized as encoding a certain type of semantic information; for example, an abstract of a story, summarizing its point, can serve as an initial boundary.

3. Formally definable internal structure. Labov has described the structure of narrative which is, in effect, a phrase-structure grammar (Labov, 1972). Planning and reasoning have been described using a transformational grammar in which the transformations mirror the real-time additions, deletions, and modifications made by speakers (Goguen, Weiner, and Linde, 1983; Linde and Goguen, 1978). A "discourse type" is a class of discourse units having internal structure in conformity with the same set of rules.

We have found that the most important discourse types in the study of crew communication are planning, reasoning, and the command-and-control speech act chain. We have also found instances of narrative and pseudonarrative, but since they are used only in non-operationally relevant ways, we do not consider them here. The following subsections discuss the three operationally relevant discourse types.

The figures show explanations of how (3) can indirectly convey (4). Figure 1 gives a list of felicity conditions for directives, a class which includes orders and requests; figure 2 gives a list of "generalizations" for the indirect accomplishment of directives. (Both figures are adapted from Searle (1979).

PREPARATORY:	ADDRESSEE IS ABLE TO PERFORM ACT A
PROPOSITIONAL CONTENT:	SPEAKER PREDICATES A FUTURE ACT A OF THE ADDRESSEE
SINCERITY:	SPEAKER WANTS THE ADDRESSEE TO DO ACT A
ESSENTIAL:	UTTERANCE COUNTS AS AN ATTEMPT BY THE SPEAKER TO GET THE ADDRESSEE TO DO ACT A

Figure 1.- Felicity conditions for directives (after Searle, "Expression and Meaning: Studies in the Theory of Speech Acts," copyrighted by and reprinted with the permission of Cambridge University Press).

1. PREPARATORY CONDITION: SPEAKER CAN MAKE AN INDIRECT DIRECTIVE TO DO ACT A EITHER BY ASKING WHETHER A PREPARATORY CONDITION CONCERNING THE ADDRESSEE'S ABILITY TO DO A HOLDS, OR BY STATING THAT IT DOES HOLD.

2. PROPOSITIONAL CONTENT: SPEAKER CAN MAKE AN INDIRECT DIRECTIVE BY ASKING WHETHER THE PROPOSITIONAL CONTENT CONDITION HOLDS OR BY STATING THAT IT DOES HOLD

3. SINCERITY CONDITION: SPEAKER CAN MAKE AN INDIRECT DIRECTIVE BY STATING THAT THE SINCERITY CONDITION HOLDS, BUT NOT BY ASKING WHETHER IT HOLDS

4. ESSENTIAL CONDITION: SPEAKER CAN MAKE AN INDIRECT DIRECTIVE TO DO AN ACT A EITHER BY STATING THAT THERE ARE GOOD OR OVERRIDING REASONS FOR DOING A, OR BY ASKING WHETHER SUCH REASONS EXIST, EXCEPT WHERE THE REASON IS THAT THE ADDRESSEE WISHES TO DO A, IN WHICH CASE THE SPEAKER CAN ONLY ASK WHETHER THE ADDRESSEE WISHES TO DO A, BUT CAN NOT ASSERT THAT HE DOES

Figure 2.- Strategies for indirect directives (after Searle, "Expression and Meaning: Studies in the Theory of Speech Acts," copyrighted by and reprinted with the permission of Cambridge University Press).

Planning- This research focusses on planning as a linguistic activity carried on by a group, rather than as an individual mental activity. The linguistic study of small group planning has shown that the language used to accomplish planning is a discourse type, since (1) it has an initial boundary, consisting of the statement of the goal which the planning is intended to accomplish; (2) it has a final boundary, which may consist of the group's evaluation of the probable effects of the plan or of their acceptance or rejection of it; and (3) it has a precise internal structure, consisting of members' proposals to add new subplans and to modify or replace parts of the plan previously proposed by others (Linde and Goguen, 1978).

Formally, the internal structure of a planning discourse unit is described as a sequence of "transformations" on the plan being formed by the group. In planning, these transformations represent the real-time effects of proposals by members to add, delete, or modify plan parts. Similarly, the relations of logical subordination that hold among the plan parts are represented by a "tree" structure. Figures 3 and 4 show a plan from the United/Portland/1978 accident. Its major goal, as stated by the first officer, is to call out the equipment (request emergency equipment be available), and his plan for this is to "have the company call." This PLAN/GOAL relationship is indicated in figure 3. In figure 4, the captain replaces the first officer's plan with a plan to "call dispatch in San Francisco." Figure 5 shows the same plan with the captain's addition of a node indicating that "maintenance down there will handle it that way."

Figure 3.- A GOAL/PLAN node.

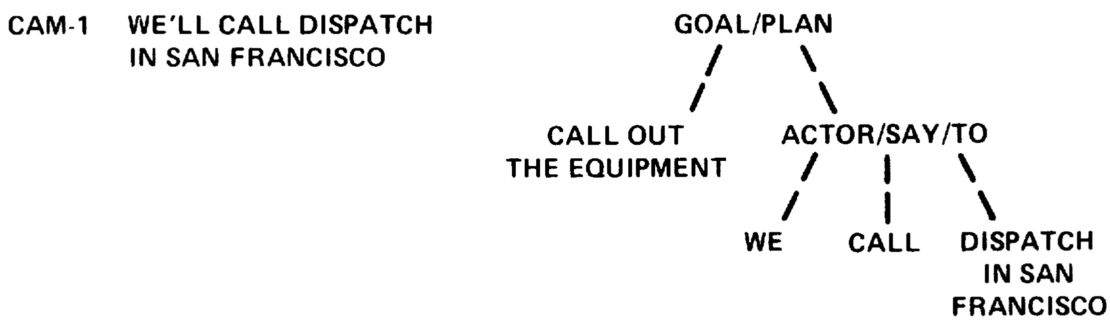


Figure 4.- Addition of an ACTOR/SAY/TO node.

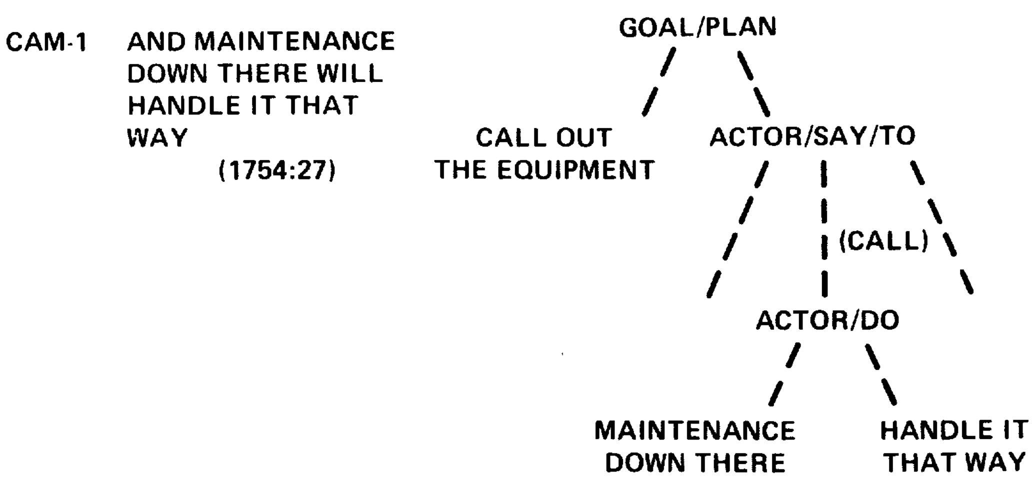
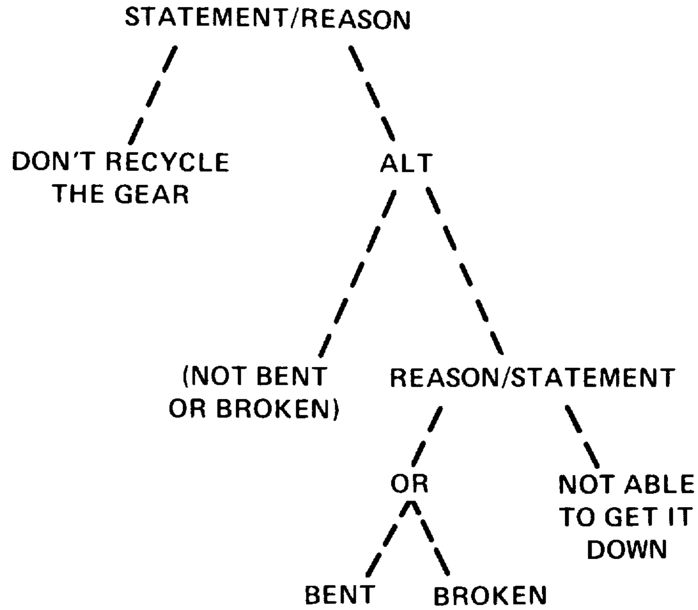


Figure 5.- Addition of an ACTOR/DO node.

Explanation- We do not use the term "explanation" to refer to segments of discourse that serve the function of explaining something; rather, explanation is a discourse type, having similar structural properties and expressible with similar formalism, as planning. Informally, an explanation is a discourse unit consisting of a statement to be demonstrated and a structure of supporting reasons, which often have further embedded relationships of subordination. This kind of discourse occurs, for example, in social contexts where a single person attempts to justify to an addressee actions he has already performed, or will perform later. Figure 6 shows an analysis of a simple explanation in which the flight engineer explains his decision not to recycle the landing gear.

The most important relationship of subordination in explanation is indicated by STATEMENT/REASON nodes. In figure 6, the main STATEMENT is "Don't recycle the



. . . AND I SAID WE'RE RELUCTANT TO RECYCLE THE GEAR FOR FEAR SOMETHING IS BENT OR BROKEN, AND WE WON'T BE ABLE TO GET IT DOWN (1751:16)

Figure 6.- An explanation tree.

gear." Everything that follows is a REASON supporting this. The ALT node represents the speaker's postulation of two alternate worlds, which differ by whether or not the landing gear is broken. This ALT node is established by the underlined portion of the text ". . . we're reluctant to recycle the gear for fear something is bent or broken." The phrase "for fear" indicates both the uncertainty about whether the gear is bent, and the decision to treat the alternate world in which it is bent as the one on which attention is focused.

Command and Control Discourse- The command-and-control speech act chain is the basic discourse type for command and control in the cockpit. A "speech act chain" is a sequence of speech acts, each having the same major propositional content. Command-and-control chains may also include the other discourse types characteristic of operationally relevant cockpit communication, planning, and explanation. The sequence of utterances (5) is a typical speech act chain. Its component speech acts include requests, reports, explanations, and acknowledgments; all concerning the major topic "fuel weight."

(5a)	CAM-1	Hey Frostie
(5b)	CAM-3	Yes sir
(5c)	CAM-1	Give us a current card on weight figure in another fifteen minutes
(5d)	CAM-3	Fifteen minutes?

(5e)	CAM-1	Yeah give us three or four thousand pounds on top of zero fuel weight
(5f)	CAM-3	Not enough
(5g)	CAM-3	Fifteen minutes is gonna really run us low on fuel here
(5h)	CAM-?	Right (1750:16)

A formal grammar specifying the constraints on command-and-control speech act chains is given in Goguen and Linde (1983). One use for such a grammar is to indicate some possible and impossible as well as optimal embeddings of social force. For example, we will not find an acknowledgment of support of a request for an act. However, we may find an acknowledgment of a request for an act and a request for support of a request for an act. We hypothesize that correct command-and-control chains describe optimal patterns of communication in the cockpit, particularly in emergency situations. Although this hypothesis cannot be tested using accident data, it could be tested with simulator experiments by training crews to use strict command-and-control form, then measuring flight performance and comparing it with performance of a control group not so trained.

VARIABLES AND ANALYTIC METHOD

This section discusses data acquisition, the theoretical concepts which define the variables used in the hypotheses, and the variables themselves.

Sampling Procedures

A sampling procedure has three main stages: (1) the production of accident transcripts, (2) the selection of transcripts, and (3) coding of the selected transcripts. The sample space that results from these procedures consists of a large number of speech acts, rather than a small number of, for example, transcripts or crew members. This choice seems well suited to studies of relationships between linguistic behavior and features of the cockpit situation. On the other hand, accident transcript data is less suitable for studies of individual differences in the behavior of crews or crew members. This is because these transcripts do not provide a sample of crews tested in a single standard situation, but rather show single crews in a variety of unique situations. We will not describe the production process of accident transcripts here, except to note that this is an "unobtrusive" procedure, in the sense that the collection of this data has nothing to do with the researchers who later analyze it.

Transcript Selection Criteria- Eight transcripts of commercial aviation accidents were selected, based on criteria using categories and analyses in (Murphy, 1980):

1. The transcript contains a critical segment. A "critical segment" is a portion of transcript containing observable degradation or failure of crew coordination which is actually or potentially critical to the completion of the flight.

2. The entire situation of interest must not be significantly longer than 30 min (since the maximum length of the tape is 30 min).

3. There must be sufficient background information to permit understanding of all relevant aspects of the situation (e.g., in the National Transportation Safety Board (NTSB) report).

4. The language of the transcript should be suitable for analysis. In particular, there should be a sufficient amount of conversation to permit analysis and all the conversation should be in English, since we are not focusing on cross-linguistic problems.

5. The probable cause factors of the accident should be sufficiently relevant to general safety issues to support further investigation.

6. All other things being equal, the most recent transcripts are preferred. (Note that this criterion also plays a major role in determining whether criterion 5 is satisfied; older flights are of lesser interest since the procedures and equipment are more likely to have been superseded.)

7. If possible, the set of transcripts should include all flight segments: taxi, takeoff, climb, cruise, approach, and land.

The transcripts selected were:

1. United Airlines/Portland/78,
2. Eastern Airlines/Miami/72,
3. Northwest Orient Airlines/Thiells, New York/74,
4. Allegheny Airlines/Rochester/78,
5. World Airlines/Cold Bay, Alaska/73,
6. Texas International Airlines/Mena, Arkansas/73,
7. Pan American Airlines/Bali/74, and
8. Air Florida/Washington, D.C./82.

<u>Data Coding</u>- Each of the 1,725 speech acts in the sample space was coded according to twelve categories: speaker, addressee(s), speech act type, discourse type, new or old topic, topic success or failure, draft order, ratification, mitigation level, crew recognized emergency, crew recognized problem, and operational relevance. Most of these variables depend upon linguistic theories that are described in the following subsection. For details of the coding procedures, see Goguen and Linde (1983).

<u>Use of Hypothesis Formulation and Test Transcripts</u>- Two of the eight transcripts, chosen for the interest of their language and situation, were used to develop hypotheses which illuminate the basic structure of crew communications. We call these two transcripts, United/Portland/78 and Texas/Mena/73, the "hypothesis

formulation group." The remaining six transcripts were used to test the hypothesis; we call these transcripts the "test group."

The six transcripts from the test group contain altogether 480 operationally relevant speech acts (defined in the "Operational Relevance" section) and the two hypothesis formulation transcripts contain 399. For each hypothesis, a dataset for testing is selected, consisting of (1) a subset of the 399 speech acts of the hypothesis formulation group, and (2) a subset of the 480 speech acts of the test group.

Each hypothesis is first tested on speech acts from the six transcripts of the test group. It is then tested on the speech acts from the two hypothesis formulation transcripts. Speech acts from these two groups are pooled when possible to yield a larger sample for a stronger test of the hypotheses. However, pooling is justified only if the methodological bias can be avoided that results from testing hypotheses on the data from which they were formulated. If the hypothesis is accepted on the basis of data from the six test transcripts and/or is rejected for data from the two hypothesis formulation transcripts, then the two datasets can be combined. The purpose of this divided testing is to reduce the probability that the obtained results are caused by an uncontrolled variable.

Variables Used in the Hypotheses

The discourse types discussed earlier provide many of the coding categories used in the precise formulation of the hypotheses, including the division of the transcript into speech acts and the inclusion of a speech act in a command and control chain or in planning or explanation. We will now discuss the remaining variables required to code the data for hypothesis testing.

Crew Recognized Emergency (CRE)- The category CRE is a social rather than a legal or factual category. The beginning of a CRE is defined as the first point at which the entire crew begins to attend to that situation which led directly to the accident. Several remarks should be made about this definition:

1. To identify the situation which led to the accident, we rely upon informed and documented opinion in the aviation community. In practice, this means that we rely on NTSB accident reports, but in disputed cases we could use minority reports, other published materials, or oral reports from members of the aviation community.

2. The definition requires that the entire crew attend to the situation. Individual crew members may have attended to the situation that led to the accident long before the CRE point and may even have attempted to bring it to the attention of the rest of the crew. However, it is group attention that is required here. Note that in practice, this implies the attention of the captain, since in the command and control situation, the captain has the authority to direct the attention of the crew to any situation which he considers to be threatening, whereas other crew members may suggest but cannot compel such attention.

3. In some accidents there may never be a CRE. These are cases in which the crew never attends to the situation that caused the accident.

Note that a captain's official Mayday declaration does not serve to establish a CRE, since this declaration often appears quite late, considerably after the point at which the crew initiates emergency actions. In fact, Mayday is a legal category, specifying a situation in which there is "immediate danger to equipment and personnel."

Crew Recognized Problem- In addition to CRE we also use the category of "Crew Recognized Problem" (CRP). This is a situation recognized by the crew as potentially dangerous and not a normal part of flight operations. It could be an actual problem or a situation which is atypical, surprising, or not expected. Like CRE, a CRP is not defined as the onset of the situation as given by the system data readout, but by the point at which the crew first attends as a whole to that situation.

Operational Relevance- A very pervasive distinction, entering into many of our definitions and all of our hypotheses, is whether a given speech act is operationally relevant. Operational relevance means that the speech act is directly involved with the achievement of the mission. Thus, a request for a snack is not operationally relevant, even though it might have some effect on the state of a crew member and hance an indirect effect on successful mission completion. This distinction has been introduced because there are certain phenomena which are potentially of great importance in operationally relevant discourse, but have no serious consequence in nonoperationally relevant segments. For example, if a speaker introduces an operationally relevant topic and other crew members do not continue this topic, the consequences can be quite serious. However, failure of a nonoperationally relevant topic is of much less concern.

Mitigation/Aggravation- The mitigation/aggravation scale developed for this study provides one way of assessing the assertiveness of speech acts. For example, (6) is direct, (7) is mitigated, (8) is highly mitigated, and (9) is aggravated.

(6) Close the window.
(7) Would you close the window?
(8) Please, would you mind closing the window?
(9) Listen, close that damn window right now.

Mitigation softens the possible offense that an utterance might give. Our results show that mitigation has a significant effect on cockpit communication, since the greater the degree of mitigation, the more likely it is that a given utterance will fail to accomplish its effect. Also, speech acts by subordinates are more mitigated than those of superiors (Section 3.1 gives a more thorough discussion of these findings).

Mitigation and aggravation are linguistic categories, not psychological ones. Thus, when a speaker uses an aggravated form, we cannot directly draw any conclusions about his psychological state at the moment, nor about his personality characteristics, although a speaker's long-term profile of using mitigation/aggravation in different contexts is probably related to his personality characteristics. The use of few mitigation strategies or of many aggravation strategies is only one way of behaving assertively. There are many devices that function to mitigate: questions are more mitigating than imperatives; modal auxiliaries such as "would," "might," and "could" are more mitigating than simple verbs; markers of request for agreement such as "right" and "OK" are mitigating. Moreover, indirect speech acts (discussed previously) are more mitigating than direct ones. (See Goguen and Linde (1983) for a review of the unified theory of mitigation given by Brown and Levinson (1979).)

Mitigation/Aggravation Scale: Several of the hypotheses proposed in this report require discrimination of degrees in a mitigation/aggravation scale as opposed to more objective measures. The degrees of this scale correspond to the feeling of native speakers of a language that some sentences are more polite (more indirect) than others. The validity of this scale has been established by checking the judgment of linguistic analysts against the judgments of members of the aviation community. We have found that four degrees of mitigation/aggravation are the most that native speakers can reliably discriminate. This scale has a midpoint of zero, representing a direct, unmitigated utterance. There are two degrees of mitigation--low and high. There is only one degree of aggravation, because aggravation is much rarer than mitigation and because there are fewer strategies for effecting aggravation than for effecting mitigation.

Scale validation was established by a reliability experiment. The subjects were asked to score 31 reports and requests chosen randomly from the six transcripts. The scale of mitigation/aggravation tested had the following four levels: Aggravated, Direct, Low Mitigation, and High Mitigation. The experimental subjects included six commercial airline professionals: two captains, three first officers, and one flight engineer. Before being asked to score the speech acts, they were given pretest training in the meaning of the categories used: a definition of mitigation was read to the subjects. They were then given written samples of speech acts to rate which were then discussed with the group by one of the researchers. Finally, they were given the written stimuli to score.

The criterion which is generally used for reliability of such scales is a stringent one: there should be at least an 80% match between the ratings of the subjects and the researchers; that is, the average number of agreements of the researchers' judgments with the subjects should exceed 8 out of 10. This criterion was just met in the present experiment, in which the average agreement of the six subjects with the researchers' judgment was 0.801. This result supports the conclusion that this is a reliable scale for degrees of mitigation. A more detailed analysis of the data (Goguen and Linde, 1983) suggests that the variance among subjects is due to regional dialect differences and to the subjects being less well trained than the researchers.

Topic and Topic Failure- A careful definition of "topic" is necessary to investigate why crew members sometimes fail to recognize or continue newly proposed topics, which are often topics of great operational importance. "Topic" is defined as the propositional content of a speech act. The "propositional content" is what a sentence predicates about the world, i.e., what the sentence is about, independent of its social force. For example, sentences (10), (11), and (12) each have different social force but the same propositional content.

(10) Close the window.
(11) The window is closed.
(12) Is the window closed?

By using this definition, we have been able to determine instances of topic failure in our sample. We define as "topic failed" any speech acts expressing a new topic not followed by a speech act from another speaker having the same topic. We have also given a taxonomy of the major topics found (see Goguen and Linde, 1983).

Draft Orders and Ratification- Plans are a major means by which a crew can discuss possible actions. A crucial question about the planning process is how decisions about taking action are actually made and expressed. This is a complex social process, requiring appropriate communications among the individuals involved and partly depending on a strict social hierarchy, in which all the participants are highly trained and are legally responsible for the correctness of the decisions made.

Any study of plan execution requires an understanding of planning as part of the command-and-control system. From the command-and-control perspective, a plan is a directive whose propositional content contains possible actions. If such a directive is made by someone other than the captain or by the captain as a suggestion rather than as an order, then it must be ratified before it has the social force of an action which the crew understands is to be performed. Since the final authority rests with the captain, all possible nonroutine actions should flow through him for ratification. Our examination of the transcripts shows that such ratifications can be either explicit or implicit. Thus, an action proposed by someone other than the captain will be called a "draft order," which requires the captain's ratification to turn it into an actual order. Actions that are proposed but not ordered by the captain are more complex; they may receive approval or modification by crewmembers and then be routed back to the captain for actual ratification. According to this plan, all ratified actions are considered to be orders issuing from the captain. These concepts are used to formulate two variables used in our hypotheses (1) whether a speech act is a draft order, and (2) whether it is subsequently ratified.

The execution of plans is relevant to aircrew coordination. A general problem here is how it can happen that important and relevant actions are not taken, e.g., an appropriate action is proposed but then not ratified.

RESULTS

The first subsection that follows describes eight research hypotheses about the use of language in this setting, reports the results of testing them, and discusses their significance for aviation safety. The next subsection summarizes the results, while the final subsection attempts to generalize the results, or discusses generalization of the results.

Hypotheses and Test Results

This subsection formulates the null hypothesis and dataset that are involved in each of our eight research hypotheses, and also gives the statistical test used and the level of significance obtained. The choice of hypotheses to be tested was influenced by the pioneering work of Foushee and Manos (1981). Each hypothesis is restricted to speech acts whose speaker and addressee are both crewmembers, because we are not studying air-to-ground communication or communication with flight attendants or passengers. They are restricted to operationally relevant speech acts because there is more linguistic variation in the nonoperationally relevant portions of the text, and because nonoperationally relevant speech acts are less important for our purpose. Flight-checklist speech acts are excluded because checklist activity is highly stereotyped; in particular, these speech acts are almost always direct and almost never acknowledged. These restrictions apply to all eight research hypotheses. A further requirement is precise definition of the variables occurring in a given hypothesis; for example, speech acts with an unknown speaker cannot be used in testing hypotheses that involve speaker rank. The frequency tables for each hypothesis are omitted here, but may be found in Goguen and Linde (1983). Other hypotheses, which we were unable to test on the present sample, may be found in Goguen and Linde (1983).

Hypothesis 1: Requests to Superiors Are More Mitigated- The null hypothesis here is that the mean mitigation/aggravation score for requests to subordinates equals (or exceeds) the mean score for requests to superiors.

The test of this hypothesis is limited to requests because requests (which include orders, questions, draft orders and suggestions) are the most characteristic speech act in command-and-control discourse and also because the consequences of misunderstanding requests are more direct and immediate than those of any other speech act. This hypothesis represents a conclusion that the speech of subordinates is more tentative and indirect than the speech of superiors. The hypothesis is important because it posits a direct effect of the basic social hierarchy on cockpit discourse. If this hypothesis is verified and if it is also shown that more highly mitigated speech acts are more often misunderstood or ignored (as is strongly suggested by the acceptance of Hypotheses 7 and 8 below), it is worthwhile to test whether training subordinates to use less mitigation would improve crew performance. Such a training hypothesis cannot itself be tested with data from accident transcripts, but could be tested with simulator experiment data.

Because the hypothesis asserts that one mean is greater than another, it is tested with a one-sided Student's t-test (t). The frequency data for this hypothesis from the six test-group transcripts yielded $t = 2.38$, degrees of freedom ($df = 136$), and probability ($p = 0.009$), using the normal approximation, which is valid because of the large sample size. The hypothesis is therefore accepted, and we conclude that the test transcript sample crewmembers indeed use more mitigation in making requests to superiors. Testing of the hypothesis with speech acts from the two hypothesis formulation transcripts yielded a similar pattern of frequencies, but with an obtained probability of only 0.32. Perhaps because there are two few speech acts to achieve the desired significance level the hypothesis is therefore not supported by these data. However, because the hypothesis has been accepted on data from the test transcripts, the speech acts from the two groups can be combined. The pooled frequencies yield $t = 2.01$ ($df = 252$, $p = 0.022$), so the hypothesis is accepted for the entire dataset.

Since appropriatness of the parametric t-test depends on homogeneity of variance, it should be noted that in this dataset, the two samples involved do indeed have approximately equal standard deviations. For speech acts from the six transcripts in the test group, the standard deviation of speech acts by subordinates is 0.516, while that of speech acts by superiors is 0.579.

<u>Hypothesis 2: Requests Are Less Mitigated in Crew Recognized Emergencies</u>- The null hypothesis here is that the mean mitigation/aggravation score for requests in CRE equals (or exceeds) the mean mitigation/aggravation score for requests not in CRE.

This research hypothesis represents our theory that when crew members know they face an emergency situation, their speech is less tentative and indirect. It is based on the notion that in any utterance, the speaker is encoding both his understanding of the situation he is talking about (the propositional content) and his understanding of the relation between himself and his addressee. Mitigation level is a major linguistic means by which a speaker can indicate his understanding of this social relation. When the situation becomes urgent, we might expect the speaker to focus most of his attention on it and thus less attention upon social relations.

Verification of this hypothesis would imply that crew members are able to vary their level of mitigation depending on their perception of the circumstances. If this is so, then training crew members to use less mitigation in specified circumstances would not seem new or strange to them, because mitigation level is already something that they alter when aware that they are in an emergency situation. If it is assumed that the actions that experienced crews take in emergency situations may be valuable, verification of this hypothesis would also lend some support to the hypothesis that training crews to speak more directly would improve their performance and thus reduce accidents (however, caution is advisable in drawing such a conclusion from the present dataset of accident transcripts).

The frequencies obtained from the test transcripts for investigating this hypothesis yield $t = 3.05$, degrees of freedom (df) = 166, $p = 0.001$, so the hypothesis is accepted. The obtained probability level for similar comparisons of speech acts in the hypothesis formulation group of transcripts is 0.026. It is therefore permissible to combine the two datasets, yielding $t = 3.46$ ($df = 276$, $p = 0.003$). Hypothesis 2 is therefore very strongly supported.

Hypothesis 3: Requests are Less Mitigated in Crew Recognized Problems- The null hypothesis here is that the mean mitigation/aggravation score for requests in CRP equals (or exceeds) the mean mitigation/aggravation score for requests not in CRP.

This research hypothesis represents a conclusion that crewmembers' speech is less tentative and indirect when they know they face a problem. Its significance is similar to that of the previous hypothesis. (Note that every CRE speech act is also a CRP speech act.)

The frequencies obtained from speech acts in the test group of transcripts, which compared CRP and non-CRP mitigation levels, were $t = 2.34$ ($df = 166$, $p = 0.010$). The hypothesis is therefore accepted for the test dataset. For the hypothesis formulation transcripts, the corresponding obtained probability level (p) is 0.149. If the two groups are combined, $t = 1.79$ ($df = 276$, $p = 0.047$). The hypothesis is therefore accepted for the dataset as a whole.

Hypothesis 4: Captains and Subordinates Differ in Frequency of Planning and Explanation- The null hypothesis is that the percentage of speech acts in explanation-and-planning discourse units produced by subordinates equals the percentage produced by superiors.

This research hypothesis indirectly addresses the effects of social hierarchy on subordinates' contributions to explaining what is happening in a situation and to planning what should happen in the future. Verification of the research hypothesis would suggest that the social hierarchy has a detrimental effect on crew communications.

We use a two-sided χ^2-test with one degree of freedom. Discourse-type frequencies for speech acts in the six test transcripts yield $\chi^2 = 1.52$ for an obtained probability level of 0.22, which does not support the research hypothesis. A similar evaluation of speech acts from the formulation-group transcripts gives $\chi^2 = 1.13$, with probability level 0.29. Therefore, the two datasets can be combined, yielding $\chi^2 = 2.97$ with $p = 0.086$. Thus the null hypothesis cannot be rejected on the pooled data and the research hypothesis is not accepted.

Modern management theory generally asserts that groups are more effective when subordinates contribute more to the group than superiors. Moreover, many informal examinations of accident transcripts have suggested that captains sometimes behave in an autocratic manner that prevents subordinates from making appropriate contributions. Based on our results, we recommend a determination of whether crew performance is improved by training subordinates to engage in more planning and explanation and also training captains to encourage this, in the condition of CRP. Furthermore, it should be determined if there are circumstances, such as CRE, in which it would be counterproductive to engage in more planning and explanation.

If we had been a priori certain of the direction of difference in frequency of planning and explaining between captains and subordinate crewmembers, we could have used a one-sided χ^2-test; then the hypothesis that captains plan and explain more would have been accepted with an obtained probability level of 0.043. Subordinates actually produced only 38% of the planning-and-explanation speech acts in the pooled dataset, whereas captains produced 62%; furthermore, captains and subordinates each produced about half of all speech acts in this dataset, but planning-and-explanation speech acts are only 9% of the total. In the future, it would be interesting to test whether captains plan and explain more during CRP and during non-CRE.

Hypothesis 5: Planning and Explanation Are Less Common in Crew Recognized Emergencies- The null hypothesis is that the percentage of speech acts that occur in planning and reasoning discourse units in a CRE equals (or exceeds) the percentage that occur in a non-CRE.

This research hypothesis represents our theory that when crew members are aware that they face an emergency, they do less planning and explaining because an emergency calls for immediate action. Precise knowledge of the distribution of planning and explanation in accident transcripts is important because it may suggest circumstances in which crews should be trained to do either more planning and explanation or less when it proves to be counterproductive.

Because the hypothesis asserts the degree of difference and there is only one degree of freedom, a one-sided χ^2-test is used. The speech act frequencies for this hypothesis in the test transcripts yield $\chi^2 = 3.87$ for an obtained probability level of 0.025. The hypothesis is therefore accepted at the 0.05 significance level. The corresponding test for speech acts from the hypothesis formulation transcripts yields $\chi^2 = 7.03$ ($p = 0.004$). Thus, it is permissible to combine the two datasets for Hypothesis 5. The combined frequencies yield $\chi^2 = 12.49$ ($p = 0.0002$); the hypothesis is therefore strongly supported on the pooled data. Further discussion of the implications of this result is included with that of the following hypothesis.

Hypothesis 6: Planning and Explanation Are More Common in Crew Recognized Problems- The null hypothesis is that the percentage of speech acts that occur in planning-and-reasoning discourse units in non-CRP equals (or exceeds) the percentage in CRP.

This research hypothesis represents the theory that crew members use more planning and explanation when they are aware that they face a problem. If verified, this hypothesis would validate the choice of the variables involved (discourse type and CRP) and would also confirm the value of training the crews to plan and reason in problem situations.

Again, we use a one-sided χ^2-test (df = 1), the discourse-type frequencies obtained from speech acts in the test transcripts yield $\chi^2 = 25.90$ with an obtained probability level beyond 0.000001. The hypothesis is therefore very strongly confirmed in this dataset. The corresponding χ^2-value for discourse-type frequencies from the hypothesis formulation transcripts is 0.27, for an obtained probability level of 0.30. Frequencies by discourse type for speech acts from the combined group of eight transcripts yield $\chi^2 = 12.03$ with a probability level of 0.0003. The hypothesis is therefore strongly confirmed for the entire dataset.

This result taken together with the findings relevant to Hypothesis 5 suggests that more planning and reasoning occur when the crew believes that it is dealing with a problem, but not when it believes that it is dealing with an emergency. One explanation for this result is that by the time an emergency situation has developed, crewmembers may feel that it is too late to take the time to plan as a group or to explain the reasons for taking specific actions. It is possible that more planning and explanation would be desirable in certain emergency situations, but not in others. To determine flight segments (if any) in which more planning and explanation produce better performance, similar experiments could be used. In any case, it is clear that crews should plan as effectively as possible during a CRP, because they may not have time for planning during a subsequent emergency.

Because this study is based upon accident transcripts, it cannot be assumed that observed crew behavior in this data is necessarily optimal. The data used in this study are probably a combination of good and bad instances of cockpit planning and reasoning; thus, testing the present hypothesis on data from normal flights should yield more definitive results.

Hypothesis 7: Topic Failed Speech Acts Are More Mitigated- The null hypothesis is that the mean mitigation/aggravation score for speech acts whose topic has failed is greater than or equal to that for speech acts whose topic has succeeded.

This hypothesis and the next one suggest that excessive mitigation can have undesirable effects in the cockpit. Since the effect of mitigation on performance (e.g., the probability of an accident) cannot be explored directly with the present data, we are forced to examine less direct connections. This hypothesis represents a conclusion that a new topic is less likely to be continued by its addressees if the speech act in which it is introduced is excessively mitigated.

Mitigation scores for the two topic conditions are compared for speech acts from the six test transcripts giving $t = 1.65$ ($df = 182$, $p = 0.01$); thus this hypothesis is accepted. For comparisons based on the hypothesis formulation transcripts, $t = 2.23$ ($df = 80$, $p = 0.013$), the combined dataset mitigation levels across topic conditions in all eight transcripts yield $t = 2.493$ ($df = 264$, $p = 0.0064$). Therefore the hypothesis is accepted.

This result lends strong support to the conclusion that excessive mitigation can have undesirable effects on crew performance. A number of NTSB reports have recommended assertiveness training for crewmembers to encourage more effective participation by subordinates. (See, for example, NTSB (1979).) Verification of the present hypothesis and the following one demonstrate effects for one kind of nonassertiveness. Moreover, this kind of nonassertiveness is defined precisely enough to allow both for training and for the evaluation of training methods.

Hypothesis 8: Unratified Draft Orders Are More Mitigated- The null hypothesis is that the mean mitigation/aggravation score for ratified draft orders equals (or exceeds) the mean for draft orders that are not ratified.

This research hypothesis tests the theory that when a crewmember proposes a suggestion to the captain, the more indirect and tentative that suggestion is, the less likely it is to be ratified by the captain. Statistical evaluation of the frequencies for ratified and unratified draft orders from the six test transcripts yields $t = 2.927$ ($df = 45$, $p = 0.002$). The hypothesis is therefore accepted for speech acts from the test transcripts. For similarly classified speech acts from the hypothesis formulation transcripts, $t = 0.589$ ($df = 13$). The t-statistic table gives an obtained probability level of approximately 0.2, so the two groups can be combined, yielding $t = 2.412$ ($df = 60$, $p = 0.008$) on the pooled data. Thus, this hypothesis is strongly supported.

As with Hypothesis 7, this hypothesis implies that excessive mitigation can have undesirable effects on crew performance. In particular, this hypothesis focuses attention on the situation in which a subordinate makes a correct suggestion that is ignored. Training in linguistic directness may be valuable in correcting this kind of pattern.

Summary of Results

This subsection includes two figures which show (1) the independent and dependent variables that are used in each hypothesis and (2) the results of testing each hypothesis.

Figure 7 shows the independent and dependent variables and which hypothesis uses each. (The two blanks suggest possibly interesting hypotheses that have not been tested in this study.) Figure 8 shows for each hypothesis: the size, N, of the dataset used to test it (in each case this includes speech acts from all eight transcripts), the obtained t-value (if any); the obtained χ^2-value; the number of degrees of freedom (for the χ^2-test), the obtained probability level for the t-test (p_t), the obtained probability level for the χ^2-test (p_χ), and the decision

	INDEPENDENT VARIABLES				
DEPENDENT VARIABLES	RANK	CRE	CRP	TOPIC FAILED	RATIFICATION
MITIGATION	1	2	3	7	8
PLAN/ EXPANATION	4	5	6		

Figure 7.- Hypotheses with dependent and independent variables.

HYPOTHESIS	N	t	χ^2	df	P_t	P_χ	DECISION
1	254	2.01	7.45	3	0.022	0.05+	YES
2	278	3.46	12.81	3	0.0003	<0.01	YES
3	278	1.79	4.70	3	0.047	<0.01	YES
4	879		2.97	1		0.086	NO
5	1039		12.49	1		0.0002	YES
6	1039		12.03	1		0.0003	YES
7	266	2.49	7.95	3	0.0064	<0.05	YES
8	62	2.41	9.52	3	0.008	0.02+	YES

Figure 8.- Summary of results.

(whether the research hypothesis was accepted). The decisions obtained using the χ^2-test agree with those obtained using the t-test, except in the case of Hypothesis 1. Although the χ^2-value is very close to that required for acceptance, it depends on the applicability of the t-test and could be rejected on that basis.

Representativeness of the Sample

We now discuss generalization of our results from these transcripts to the broader population of commercial-aviation cockpit discourse. The results will generalize if the sample is representative. This subsection argues for the representativeness of our sample.

It might be argued that the sample cannot be representative because it consists of only eight transcripts. But the sample consists not of the eight transcripts (or equivalently, the eight crews) nor of the 25 speakers quoted in the transcripts, but rather of all the operationally relevant speech acts produced by these speakers. This is a much larger sample which is much more likely to be representative of its population, for the following reasons given.

The basic criterion for representativeness is that a sample is very likely to be representative if it is sufficiently large and is also a random sample; in fact, the probability that a random sample is not representative can be made as small as desired by making the sample large enough. Our experience with statistical studies of other linguistic data suggests that samples of one or two hundred units are generally adequate (Herdan, 1966) and that smaller samples often suffice if the pattern of variation is not especially complex (Guy, 1980). Only Hypothesis 8 might be considered nonrepresentative because of sample size; however, as this hypothesis is not precise, the small sample size should simply be noted.

An argument for representativeness is that the sample can be successfully used as a standard of comparison for the behavior of crewmembers, that is, the sample is homogenous. Since we have data collected from a number of speakers, it might be questioned whether the sample is dominated by a few loquacious speakers who exhibit unusual linguistic behavior. To verify the assertion that individual differences are relatively unimportant in this sample compared to differences arising out of the cockpit situation in which the language is produced, we tested whether the most loquacious speaker of each rank differs significantly from his colleagues of the same rank, using the degree of mitigation/aggravation scale. We found that the sample was not dominated by a few speakers with unusual linguistic behavior.

We now give three arguments for the randomness of our sample. The first and most direct argument is that our sample is random because the criteria used for transcript selection are statistically independent of the dependent measures used in the hypotheses. For example, the occurrence or nonoccurrence of a critical segment in a transcript cannot affect the mitigation level of the speech acts occurring in that transcript. (The Transcript Selection Criteria section gives the selection criteria used in this study.)

The second argument is based on the principle of locality of effects, which states that although there are significant sequential dependencies in natural language, they are largely confined to units a few steps earlier in the sequence and hence have little effect on the randomness of any reasonably large sample. This principle has been observed for units at all levels of the linguistic hierarchy, including phonemes, morphemes, lexemes (i.e., words), and syntactic phrases. This principle is presumed to hold for speech acts as well, although this assumption has not been formally tested.

For the third argument, recall from the section Use of Hypothesis Formulation and Test Transcripts that the transcripts are divided into two disjoint groups, called the hypothesis formulation and test groups, to reduce the likelihood that the result obtained from testing a given hypothesis is due to an uncontrolled variable which is different from the independent variable of that hypothesis.

The results of the statistical tests on the research hypotheses of this study are valid as descriptive statistics, that is, as statistical summaries of a particular sample. The above arguments for extent of generalization support our giving to these results the usual inferential interpretation.

CONCLUSIONS AND EXTENSIONS

The results reported in this study support the utility of a methodology for the detailed analysis of cockpit discourse, and its application to improving aviation safety. This methodology has produced verified results concerning the linguistic behavior of aircrews that have significant implications for crew training. In addition, possibilities for further research have been suggested. The first two subsections below detail the main contributions of this research and the third discusses its extensions.

Basic Contributions

1. A theory of the structure of command-and-control discourse that includes a determination of its relationships to planning and explanation, as well as determination of its basic speech acts, which are request, report, acknowledgment, and declaration.

2. A general theory of the structure of discourse; this theory involves analyzing a given discourse unit as a sequence of transformations that have an underlying tree structure representing the structure of the discourse, i.e., a hierarchical classification of the discourse parts and their relationships.

3. A classification of the discourse types that occur in aviation discourse. These are: command-and-control speech-act chain, including the subtype of checklist; planning; explanation; and narrative and pseudonarrative.

4. A scale of mitigation levels for speech acts occurring in aviation discourse. This five-point scale ranges from "highly mitigated" to "aggravated" and has "direct" as its zero point. The scale has been experimentally validated.

5. A theory of draft orders and the process of ratification has been developed, based on the theories of planning, explanation, and command-and-control discourse.

6. A collection of variables that summarize many important characteristics of the speech acts that occur in cockpit discourse.

Applied and Specific Contributions

This subsection relates the hypotheses that have been discussed to further research and direct training in crew coordination and communication for aviation safety. These results are necessarily limited by the restriction of our data to accident transcripts. It should be possible to go much further in the directions indicated here whenever both systems data and nonaccident data are available.

1. The average mitigation level of requests by subordinates is significantly higher than that of requests by superiors. Testing should be done on whether this asymmetry contributes to the misinterpretation of suggestions and commands in the cockpit, because it should not be difficult to train subordinate crewmembers to use less mitigated language, termed "to be more assertive" in the NTSB reports.

2. Requests are less mitigated during a CRP and are still less mitigated during a CRE. This suggests that crewmembers should not find it strange or abnormal to be trained to use less mitigation, since variation of mitigation level is something that they already use under certain conditions. It also suggests that training in linguistic assertiveness would only be reinforcing a tendency that already appears in problem and emergency conditions.

3. Superiors produce a higher proportion of explanation-or-planning speech acts than subordinates. The optimal ratio is not clear and may depend on the context; this should be investigated. There are reasons to believe that this ratio would be a good indicator of degree of the authority delegated by a given captain to his crew.

4. Planning and explanation occur much more frequently during CRP, and that they are less common during CRE. This suggests further research to discover whether training crewmembers to engage in more planning and reasoning under real emergency conditions would improve performance.

5. The more-mitigated speech acts which introduce a new topic are less likely to have their topic become the subject of further conversation. This demonstrates the importance of crew members not using mitigated language when introducing operationally significant topics. Because this also is behavior for which crewmembers presumably can be trained, we suggest further exploration of both of the basic linguisitc phenomena, testing to determine whether such training can improve any objective performance measures.

6. On the average, draft orders that do not get ratified are more mitigated than those that do get ratified. This was shown with a very high level of significance. The implications of this result are very similar to those of the previous result, but concern the ratification of subordinates' suggestions rather than the success of their topics.

7. A number of other linguistic variables should be investigated for correlation with objective-system and crew-performance variables (see Goguen and Linde, 1983). These variables include degree of command-and-control coherence, the rate of request-report-acknowledge triples, the rate of planning and reasoning, and the rate of simple acknowledgments. In certain cases, it might be less costly to use a reliable linguistic variable as an indicator of some objective performance measure than to measure it directly. In other cases, important training implications might be discovered.

8. Finally, the research program initiated in this report should have many applications to the design of aviation procedures and equipment systems that involve communication (such as on-board display and speech generation). This possibility of application arises from the clear demonstration that air-crew discourse involves definite linguistic structures, and that these structures correspond in specific ways to the operational structure of the flight. This means that there are only certain times when it is natural for certain kinds of communications to occur, and that there are natural forms for each kind of communication. For example, a piece

of equipment in the cockpit that produced complex verbal information about the status of the flight plan would probably not be useful unless it produced this information at the right time and in the right form.

Extensions

The methodology described here is presently being used in a study of crew coordination factors and their relationship to flight task performance, using as data the audio and video recordings of 16 full-mission simulations (Murphy et al., 1984). This will permit some important extensions of the present research including (1) extension of the hypotheses to nonverbal performance and to factors of the aircraft systems, and (2) comparison of communication in successful and unsuccessful flight performance. We expect that the findings of the current study will be confirmed, refined and extended by this richer dataset. We hope this will lead to the development of new methods for training crews in more effective communication and will provide guidelines for the design of improved aviation procedures and equipment.

REFERENCES

Brown, P.; and Levinson, S.: Universals in Language: Politeness Phenomena. Esther N. Goody, ed., Questions and Politeness: Strategies in Social Interaction. Cambridge University Press, 1979.

Foushee, H. C.; and Manos, K. L.: Information Transfer within the Cockpit: Problems in Intracockpit Communications. NASA Technical Paper 1875, Information Transfer Problems in the Aviation System, C. E. Billings and E. S. Cheaney, eds., 1981, pp. 63-71.

Gazdar, G.: Pragmatics: Implicature, Presupposition and Logical Form. Academic Press, 1979.

Goguen, J.; and Linde, C.: Linguistic Methodology for the Analysis of Aviation Accidents. Technical Report, Structural Semantics, 1983, and NASA CR-3741, 1983.

Goguen, J.; Weiner, J.; and Linde, C.: Reasoning and Natural Explanation. Int. J. Man-Machine Studies, vol. 19, 1983, pp. 521-599.

Gordon, D.; and Lakoff, G.: Conversational Postulates. Papers from the Regional Meeting, Chicago Linguistics Society, 1971.

Guy, G.: Variation in the Group and the Individual: The Case of Final Stop Deletion. William Labov, ed., Locating Language in Time and Space. Academic Press, 1980.

Herdan, G.: The Advanced Theory of Language as Choice and Chance. Springer-Verlag, 1966.

Labov, W.: The Transformation of Experience into Narrative Syntax. Language in the Inner City. University of Pennsylvania Press, Philadelphia, 1972.

Labov, W.; and Fanshel, D.: Therapeutic Discourse; Psychotherapy as Conversation. Academic Press, 1977.

Linde, C.: The Linguistic Encoding of Spatial Information. Ph.D. thesis, Columbia University, 1974.

Linde, C.; and Goguen, J.: The Structure of Planning Discourse. J. Social and Biological Structures, vol. 1, 1978, pp. 219-251.

Linde, C.; and Labov, W.: Spatial Networks as a Site for the Study of Language and Thought. Language, vol. 51, 1975.

Murphy, M.: Analysis of Eighty-four Commercial Aviation Incidents: Implications for a Resource Management Approach to Crew Training. IEEE Proceedings, Annual Reliability and Maintainability Symposium, 1980.

Murphy, M.; Randle, R.; Tanner, T.; Frankel, R.; Goguen, J.; and Linde, C.: A Full Mission Simulator Study of Aircrew Performance: The Measurement of Crew Coordination Factors and Their Relation to Flight Task Performance. Hartzell, E. J. and Hart, S., eds., Papers from the 20th Annual Conference on Manual Control. NASA CP-2341, 1984.

National Transportation Safety Board: Aircraft Accident Report--United Airlines, Inc., McDonnell-Douglas DC-8-61, N8082U, Portland. Technical Report, National Transportation Safety Board, 1979.

Sachs, H.: An Analysis of the Course of a Joke's Telling in Conversation. Bauman, R. and Sherzer, J., eds., Explorations in the Ethnography of Speaking. Cambridge University Press, 1974.

Searle, J.: Expression and Meaning. Cambridge University Press, 1979.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM 88254		

4. Title and Subtitle	5. Report Date August 1986
CREW COMMUNICATION AS A FACTOR IN AVIATION ACCIDENTS	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Joseph Goguen,* Charlotte Linde,† and Miles Murphy	A-86254
	10. Work Unit No.
9. Performing Organization Name and Address Ames Research Center, Moffett Field, CA 94035 *Structural Semantics, Palo Alto, and SRI International, Menlo Park, CA. †Structural Semantics	11. Contract or Grant No.
	13. Type of Report and Period Covered Technical Memorandum
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546	14. Sponsoring Agency Code 505-35-21

15. Supplementary Notes

Point of contact: Miles Murphy, Ames Research Center, MS 239-3, Moffett Field, CA 94035, (415)694-5906 or FTS 464-5906

16. Abstract A method for the detailed analysis of within-crew communications is developed and applied in formulating and testing several hypotheses about the basic structure of the aircrew communication process. Planning and explanation are shown to be well-structured discourse types, described by formal rules. These formal rules are integrated with those describing the other most important discourse type within the cockpit: the command-and-control speech act chain. Command-and-control discourse is described as a sequence of speech acts for making requests (including orders and suggestions), for making reports, for supporting or challenging statements, and for acknowledging previous speech acts. Mitigation level, a linguistic indication of indirectness and tentativeness in speech, was an important variable in several hypotheses. Testing these hypotheses showed that the speech of subordinates is more mitigated than the speech of superiors, that the speech of all crewmembers is less mitigated when they know that they are in either a problem or emergency situation, and that mitigation is a factor in failures of crewmembers to initiate discussion of new topics or have suggestions ratified by the captain. The test results also indicated that planning and explanation are more frequently performed by captains than by other crewmembers, are done more during crew-recognized problems, and are done less during crew-recognized emergencies.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Aviation accidents, Social psychology, Group processes, Decision making, Communication, Linguistics	Unlimited Subject category - 53

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	26	A03

*For sale by the National Technical Information Service, Springfield, Virginia 22161